



Munich Personal RePEc Archive

Risk-based decisions on assets structure of a bank — partially observed economic conditions

Hałaj, Grzegorz

06. August 2006

Online at <http://mpra.ub.uni-muenchen.de/523/>

MPRA Paper No. 523, posted 07. November 2007 / 01:06

Risk-Based Decisions on Assets Structure of a Bank — Partially Observed Economic Conditions.

Running title: Asset structure with partial observation.

Grzegorz Hałaj*

August 6, 2006

ABSTRACT

A model of bank's dynamic asset management problem in case of partially observed future economic conditions and requirements concerning level of risk taken has been built. It requires solving the resulting optimal control with random terminal condition resulting from partial observation of parameter of maximized functional. Stochastic Maximum Principle reduces the problem to solving FBSDE. As optimization may usually imply dependence of forward equation on solutions of backward equation we allow the drift and diffusion of forward part to be functions of solution of backward equation. The necessary conditions for existence of solutions of FBSDE in such a form have been derived. A numerical scheme is then implemented for a particular choice of parameters of the problem.

Keywords: Portfolio optimization; bank's assets; partial observation; stochastic maximum principle;+ FBSDEs.

*National Bank of Poland, Świętokrzyska 11/21, 00-919 Warsaw, e-mail: grzegorz.halaj@n-s.pl. The author is a PhD student of Prof Łukasz Stettner from Institute of Mathematics Polish Academy of Sciences.

1 Introduction

We aim to solve an optimal asset management problem of a bank over the finite horizon. It can be a problem of management of the bank but it can be a problem of banking supervision that want to forecast activity of banks. Nevertheless, in both cases, there are some conditions of banking environment that have to be taken into account in asset structure optimization. Every bank is subject to certain limitations considering risk taking (Basel II, 1999). For example it is required that bank holds sufficient amount of capital to cover expected losses. It must report to banking supervision the so called Capital Adequacy Ratio (CAR). If CAR is less then the well established threshold value of 8% then bank may face severe consequences resulting in the extreme case in withdrawal of banking licence. The strict instructions how CAR has to be calculated are described in each country banking legislation. However, there are some standards of how soundness of banks' capital is to be measures gathered in Capital Adequacy Accord prepared by Bank of International Settlements. The ratio depends strongly on the structure of assets possessed by the bank and this feature will be reflected in the mathematical formulation of the problem. In a very general terms it can be said that the more cash and government securities the bank holds the higher is CAR and the more mortgage or investment loans are granted the lower is the adequacy ratio.

However, the simple rule to retain CAR below 8% threshold by investing only in T-bills and concentrating on secure lending may be detrimental to banks profitability. Bank has to take risk to have the opportunity to increase its wealth, to be able to pay dividends and to build its capital base.

What makes the decisions of the bank more difficult, usually *future mar-*

ket conditions are uncertain. For the purpose of the article, future market conditions are defined as the overall economic situation on the market after the given investment horizon. They are usually unknown or at least the recognition of them in longer perspective is very difficult, but they can be observed through market indicators. For example, the level of inflation can quite accurately be predicted for a month in advance. Its level after a year or 5 years is very uncertain and the market can learn about it as the time elapses. Why this future conditions may matter in bank's decision making within finite period of time? Let's look on the problem from 2 perspectives.

1. (*bank's perspective*) Let's assume that after that period bank invests and consumes according to the solution of classical Merton optimization problem with asset return given by mean predicted economic conditions. If the internal models of the bank indicate favorable economic outlook in the later future but currently the return from stocks or loans to individuals is low then it might be optimal for a bank to invest firstly in a secure way to minimize the risk of default and to probably show good results in the next (future) period.
2. (*perspective of supervision authority*) Banking supervision may be interested in predicting the strategy of a bank given model applied by the bank to manage its portfolio. If it can be shown that models tend to overestimate the situation of the market it might be useful to formulate preemptive instructions for the bank to prevent it from increasing of the riskiness of assets because it may require additional capital. It may happen like that if the market condition is not as good as models tend to indicate.

We make here a very important assumption. The asset prices depend on the *future market conditions* only through market beliefs about those conditions. Usually prices reflect important macro- as well as microeconomic factors but in the end there are market participants' buying/selling decisions that influence the moves of stock prices. Hence, in models assuming efficiency of markets the decision are based on the observed prices only. But in case of a huge, sophisticated institutions like banks it may be useful to consider its own and maybe unique perception of risk factors. It may influence its decisions. The bank builds complicated internal model to assess market conditions. The models may on one hand let them to more deeply investigate the economic factors and business perspective. In this way they may for instance want to work out the strategy defending them against economic slowdown and helping to retain sufficient amount of capital for the future. On the other hand they may incorporate beliefs of managers about the economic conditions.

We have built a model of bank's decision making that takes into account the four essential elements of banking environment: risk, profitability and bank's as well as market predictions about uncertain *future market conditions*.

There are many examples of optimization problems with partial observation and their application to modeling economic problems, in particular optimal investment and consumption issues. Linear problem solutions can be characterized by means of Kalman-Bucy filters — in general case for example like in Bensoussan (1992); Lipster and Shiryaev (2001); in financial case like in Brendle (2005); Lefèvre (2001); Pastor and Veronesi (2003). Unlike in linear problems, nonlinear filtering leads usually to complex infinite

dimensional filters, e.g. described by means of Zakai partial SDEs (Benès et al., 2004; Carmona and Ludkovsky, 2004). On the other hand, there is a range of papers dealing with optimization of portfolios with limited risk and developing theory of imposing risk regulation in banks, e.g. see. Cuoco and Liu (2005); Emmer et al. (2001); Santos (2002).

In our model we use the special case of terminal wealth maximization with the penalizing cost of investing in too risky or in too secure way to describe the bank's optimization problem. The utility of terminal wealth is weighted by future economic situation that is partially observed, conditioned on the information available on the market. Now we know that randomness in goal function may appear if we consider partial observation problem. The randomness in the goal function related to partial observation may cause obstacles if we want to use dynamic programming technics to obtain optimal controls. In such a case maximum principle approach proved to be helpful. However, in this case we are facing problems with obtaining the explicit solution. We have to use numerical methods with good convergence properties. We will follow the reasoning of Zhang (2004) in a numerical analysis of the problem. Similar methods have been studied by Rivière (2005) or Delarue and Menozzi (2005) and near-optimality by Zhou (1998).

The paper has the following structure. Firstly, we present formal model of the market and decision making. Secondly, we formulate stochastic maximum principle suitable for our problem. Thirdly, we give conditions for solvability of related adjoint equations. Finally, we solve the numerical example and discuss convergence of investment policy obtained by means of numerical scheme to a theoretical one. The most frequently used notations

are presented below¹.

- \mathbb{H} – filtration $\{H_t\}_{t \in [0, T]}$;
- $\Lambda_T^p(\mathbb{H})$ — the space of \mathbb{H} -predictable processes $\{X(t)\}_{t \in [0, T]}$, satisfying for every $t \in [0, T]$ almost surely $\int_0^t |X(s)|^p ds < \infty$;
- $\Lambda_T^p(\mathbb{H})$ — for RHSC filtration \mathbb{H} ($\mathcal{H}_T \subset \mathcal{F}$), the space \mathbb{H} -adaptable processes $\{X(t)\}_{t \in [0, T]}$, satisfying for every $t \in [0, T]$ almost surely the condition $\int_0^t |X(s)|^p ds < \infty$;
- $L_{\mathcal{F}}^p(\mathbb{R}^n)$ — for $p \in (0, \infty)$ the space of functions with the domain in \mathbb{R}^n , \mathcal{F} -measurable, such that for $X \in L^p(\mathbb{R}^n)$ we have: $\mathbf{E}|X|^p < \infty$; $L_{\mathcal{F}}^\infty(\mathbb{R}^n)$ — space of \mathcal{F} -measurable functions X such that $\sup_{\omega \in \Omega} |X(\omega)| < \infty$;
- $\mathcal{L}_T^2(\mathbb{R}^{n \times m}, \mathbb{H})$ — the space of processes $\{X(t)\}_{t \in [0, T]}$ n -dimensional, \mathbb{H} -adaptable, satisfying the condition $\mathbf{E} \int_0^T |X(t)|^2 dt < \infty$; in this case $|X(t)| = \left(\sum_{i=1}^n \sum_{j=1}^m |X_{ij}(t)|^2 \right)^{\frac{1}{2}}$
- $\mathcal{L}_T^\infty(\mathbb{R}^n, \mathbb{H})$ — the space of processes $\{X(t)\}_{t \in [0, T]}$ n -dimensional, satisfying condition $\mathbf{E} \max_{0 \leq t \leq T} |X(t)|^2 dt \leq \infty$.

2 Problem to be solved

2.1 Mitigation of the risk of return

Let's concentrate on formalization of bank decision problems. Usually the goal of a bank is to generate high results in the least risky way that is pos-

¹We will be omitting brackets if it is not confusing — e.g. Λ_T^p instead of $\Lambda_T^p(\mathbb{H})$ if it is clear which σ -field we have meant.

sible. High return and low risk are two contradictory goals and a bank faces a problem of choosing such an investment policy so as not to take too risky exposure but guarantee a satisfactory return. For example a bank's management can ask how to expend lending activity without a need to inflate too much capital requirement. The bank can simply be afraid of dropping below 8% capital adequacy requirement what in turn would trigger supervisory actions. The bank's problem could be translated in the general terms to the following optimal control.

Investment horizon. Bank plans its activity within a finite horizon T . Why may the bank be interested in planning in a finite horizon? It makes sense to concentrate the attention on particular finite period of time since bank may want to or have to show results to investors or current shareholders after that time. If there is market potential for high profitability, showing high return on assets (ROA) or return on equity (ROE) may attract investors and broaden sources of funding and may lower cost of initial public offering (IPO) if needed.

Economic situations. The risk in the model is introduced by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ satisfying standard assumptions. We introduce two independent, 1-dimensional (\mathbb{F}, \mathbf{P}) -Brownian motions $W^{(1)}$ and $W^{(2)}$. The market savvy enables the bank to formulate and verify the hypothesis that given current market situation the prediction of market conditions satisfy the following equation: $d\lambda(t) = \gamma(t, \lambda(t))dt + \sigma_\lambda^{(1)}dW^{(1)}(t)$, γ — deterministic and Borel function. However, the bank can only estimate λ from the observed process $\hat{\lambda}$. The bank treats this predictions as binding in the decision process. We assume that market participants believe that economic

situation can be described by this factor $\{\hat{\lambda}(t)\}_{t \in [0, T]}$. It is assumed to satisfy a diffusion equation

$$d\hat{\lambda}(t) = \hat{\gamma}(t, \lambda(t), \hat{\lambda}(t))dt + \sigma_{\lambda}^{(2)}dW^{(2)}(t)$$

with a deterministic, Borel function $\hat{\gamma}$. A parameter $\sigma_{\lambda}^{(2)}$ is a deterministic rate of bank's inaccuracy in assessing true economic conditions λ . However its dynamics is not known by the market participants and only asset prices are influenced directly by observed $\hat{\lambda}$. We assume that $W^{(1)}$ and $W^{(2)}$ are independent.

Market. Let's assume the following asset's market $\{S_i: i \in \{0, 1, \dots, n\}\}$ where bank operates. The process $S_0(t) = e^{rt}$ denotes the price dynamics of risk-free zero-coupon bond for an interest rate $r \geq 0$ and S_i be the price of risky securities which satisfy $\forall i \in \{1, \dots, n\}$ the following equation:

$$dS_i(t) = S_i(t) \left[(\mu_i(t) + m(\hat{\lambda}(t)))dt + \sum_{j=1}^d \sigma_{ij}(t)dW_j(t) \right], \quad (1)$$

where $\mu_i(t)$ is deterministic mean rate of return from the investment in risky assets and $m(\hat{\lambda}(t))$ is \mathbf{P} -a.s. bounded process. A number $\sigma_{ij}(t)$ is risk scaling parameter telling us how much the price of i th security is volatile and differs in the future from expected value as a consequence of the presence of risk factor W_j — Brownian motion adapted to \mathbb{F} and s.t. for all $i \neq j$ W_i is independent from W_j . We denote

$$W: = [W_1^\top \dots W_n^\top].$$

μ can be treated as the historic estimation of return and $m(\cdot)$ reflects the correction of historical return by predictions of *future market return*. We assume that W is independent of $W^{(1)}$ and $W^{(2)}$.

We thus introduce the filtration $\mathbb{G}: = \{\mathcal{G}_t | t \in [0, T]\}$ with $\mathcal{G}_t: = \sigma\{S_u, \hat{\lambda}_u | u \leq t\}$ available to the bank.

Investment decision. The wealth process $X^{(\pi)}$ — when the bank invests $\pi_i(t)$ units of money from its wealth of into asset i at time t — has the following differential representation:

$$dX^{(\pi)}(t) = rX^{(\pi)}(t)dt + (\mu(t) + m(\hat{\lambda}(t)) - r\mathbf{1}_n)\pi(t)dt + \pi(t)\sigma(t)dW(t). \quad (2)$$

The risk manager is supposed to adapt investment policy π maximizing

$$J(\pi) = \mathbf{E} \left[\int_0^T -e^{\delta t} C(t, \pi(t)) dt + e^{\delta T} \mathbf{E} [U(\lambda(T)X^{(\pi)}(T)) | \mathcal{G}_T] \right] \quad (3)$$

where C is the cost of applying policy $\pi(t)$ which is to risky and may result in huge losses or can easily lead to exceeding of certain legal limitation (e.g. fall of CAR below 8% level). Parameter δ is a constant discount factor. The set of admissible controls is denoted by \mathcal{A} and is given in the following way:

$$\mathcal{A}: = \{\pi \in \mathcal{L}_T^2(\mathbb{R}^n, \mathbb{G})\}. \quad (4)$$

The optimal policy in \mathcal{A} will be denoted by π^* , i.e. $\pi^*: = \arg \max_{\pi \in \mathcal{A}} J(\pi)$ (optimal wealth process $X^{(\pi^*)}$ — by X^*). Dependence of cost function on time t can be used to model changing in time benchmark of cost (e.g. *accepted level of risk*). The manager may assume a given level of asset volatility and then try to optimize future return $X^{(\pi)}$. The parameter $\hat{\lambda}$ has an interpretation as in Section 1. The function o should be increasing to model better returns from investing $X^{(\pi)}(T)$ in the period $[T, +\infty)$ in very good economic situation after T and losses (in extremely) bad market environment (e.g. huge credit risk factor — the bank for which $X^{(\pi)}$ serves as a model of credit risk exposure receives only $\lambda(T)X^{(\pi)}(T)$, or $\lambda(T)$ fraction of loan portfolio receivables).

3 Stochastic Maximum Principle

Change of measure. We would like to transform measure \mathbf{P} to such a measure \mathbf{P}_0 that \mathbb{G} is Brownian filtration with respect to \mathbf{P}_0 . Than for any random variable V that is \mathcal{F} -measurable $\{\mathbf{E}_0[V|\mathcal{G}_t]\}_{t \in [0, T]}$ would have nice integral representation. For the process

$$\Lambda(t) := \exp \left(\int_0^t \frac{\hat{\gamma}(s, \lambda(s), \hat{\lambda}(s))}{\sigma_\lambda^{(2)}} dW_\lambda^{(2)}(s) - \frac{1}{2} \int_0^t \left(\frac{\hat{\gamma}(s, \lambda(s), \hat{\lambda}(s))}{\sigma_\lambda^{(2)}} \right)^2 ds \right)$$

we assume that it is a (\mathbf{P}, \mathbb{F}) martingale. We define a measure \mathbf{P}_0 on \mathcal{F} with the Radon-Nikodym derivative $\frac{d\mathbf{P}_0}{d\mathbf{P}} = \Lambda(T)$. Then $\{\hat{\lambda}(t)\}_{t \in [0, T]}$ is $(\mathbf{P}_0, \mathbb{G})$ -Brownian motion. What is more, $\hat{\lambda}$ and λ are \mathbf{P}_0 -independent and the process

$$\bar{W}(t) = W_\lambda^{(2)}(t) + \int_0^t \frac{[\hat{\gamma}(s) - \mathbf{E}[\hat{\gamma}(s) | \mathcal{G}_s]]}{\sigma_\lambda^{(2)}} ds$$

is innovation process, i.e. a (\mathbf{P}, \mathbb{G}) Brownian motion.

Stochastic Maximum Principle, taken from Yong and Zhou (1999), incorporates the risk-adjusted control into Hamiltonian system. This is the essential difference between deterministic and stochastic case. But we depart from considering *such a* general case. We take a region \mathbf{U} in a given metric space. We need a technical notion of *extended Lipschitz continuity* for existence of solution to controlled SDE (Yong and Zhou, 1999).

Definition 3.1 *Measurable function $\phi : [0, T] \times \mathbb{R}^n \times \mathbf{U} \rightarrow \mathbb{R}^n$ satisfies Lipschitz condition if and only if there exists $L > 0$ and i.e. modulus of continuity $\kappa : [0, \infty) \rightarrow [0, \infty)$ (i.e. the monotonic and continuous function with $0 \leftrightarrow 0$) such that $\forall t \in [0, T], x, \tilde{x} \in \mathbb{R}^n, u, \tilde{u} \in \mathbf{U}$ it is true that:*

$$1. \quad |\phi(t, x, u) - \phi(t, \tilde{x}, \tilde{u})| \leq L|x - \tilde{x}| + \kappa(d(u, \tilde{u})),$$

$$2. \quad |\phi(t, 0, u)| \leq L.$$

We consider a state process $X^{(u)}$ satisfying for a given \mathbf{U} -valued process u the equation

$$\begin{cases} dX^{(u)}(t) &= b(t, X^{(u)}(t), \hat{\lambda}(t), u(t))dt + \Sigma(t, X^{(u)}(t), u(t))dW(t) \\ X^{(u)}(0) &= x_0 \in \mathbb{R}^k, \end{cases} \quad (5)$$

for b, σ satisfying extended Lipschitz condition. Then (5) admits a unique solution (Yong and Zhou, 1999, Chapter 1, Section 6.4, theorem 6.17). For the functions

$$f: [0, T] \times \mathbb{R}^k \times \mathbf{U} \rightarrow \mathbb{R} \quad (6)$$

$$h: \mathbb{R}^k \rightarrow \mathbb{R} \quad (7)$$

satisfying enhanced Lipschitz condition and bounded Borel function $o: \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ a functional $J: \mathcal{A} \rightarrow \mathbb{R}$ defined for the family $\mathcal{A} = \{u: [0, T] \times \Omega | u \text{ is } \{F_t\}_{t \in [0, T]} \text{-adapted, } \{\sigma(t, X^{(u)}(t), u(t))\}_{t \in [0, T]} \in \mathcal{L}_T^2(\mathbb{R}^{k \times n}, \mathbb{G})\}$ is described by the following expression:

$$J(u) = \mathbf{E} \left[\int_0^T f(t, X^{(u)}(t), u(t))dt + \mathbf{E} \left[h(o(\lambda(T))X^{(u)}(T)) \middle| \mathcal{G}_T \right] \right]. \quad (8)$$

The optimization problem can be stated as follows:

Problem 3.1 *Find a process $\bar{u} \in \mathcal{A}$ satisfying:*

$$J(\bar{u}) = \sup_{u \in \mathcal{A}} \{J(u)\}. \quad (9)$$

A pair consisting of the process \bar{u} satisfying (9) and the corresponding state process $X^{(\bar{u})}$ will be called a solution to the *functional maximization control (FMC)*.

To formulate maximum principle we need two core notions — *Hamiltonian* function and *adjoin equation* to equation (5).

Definition 3.2 *The function $H: \Omega \times [0, T] \times \mathbb{R}^k \times \mathbf{U} \times \mathbb{R}^k \times \mathbb{R}^{k \times n}$ such that:*

$$H(\omega, t, x, u, p, q) = p^\top b(t, x, \hat{\lambda}(\omega, t), u) + \text{tr}(q^\top \Sigma(t, x, u)) + f(t, x, u) \quad (10)$$

is called the Hamiltonian function (for Problem (3.1)). We will drop ω in H .

Definition 3.3 *The adjoin equation to Equation (5) is Backward Stochastic Differential Equation (BSDE):*

$$\begin{cases} dp(t) &= -H_x(t, X^{(u)}(t), u(t), p(t), q(t))dt + q(t)dW(t) + q_\lambda(t)d\hat{\lambda}(t) \\ p(T) &= \mathbf{E} \left[o(\lambda(T))h_x(o(\lambda(T))X^{(u)}(T)) \mid \mathcal{G}_T \right] \end{cases} \quad (11)$$

solved for a triple $(p, q, q_\lambda) \in \mathcal{L}_T^2(\mathbb{R}^k) \times \mathcal{L}_T^2(\mathbb{R}^{k \times n}) \times \mathcal{L}_T^2(\mathbb{R})$ such that $\forall t \in [0, T]$

$$\begin{aligned} p(t) &= \mathbf{E} \left[o(\lambda(T))h_x(o(\lambda(T))X^{(u)}(T)) \mid \mathcal{G}_t \right] + \\ &\quad + \mathbf{E} \left[\int_t^T H_x(s, X^{(u)}(s), u(s), p(s), q(s))ds \mid \mathcal{G}_t \right]. \end{aligned}$$

The component that tells the stochastic and deterministic maximum principle apart is the additional adjoin equation for $p(t)$. Stochastic maximum principle (containing sufficient condition) that has been cited is a special case of more general one in the sense that the diffusion coefficient is differentiable with respect to $u(t)$. Therefore, we are getting rid of the problem of incorporating additional risk-adjusted adjoin equation.

Theorem 3.1 (Stochastic maximum principle) *Assume that for fixed t, p, q the functions $H^{(t,p,q)}: \mathbb{R}^k \times \mathbf{U} \rightarrow \mathbb{R}$ given by*

$$H^{(t,p,q)}(x, u) = H(t, x, u, p, q) \quad (12)$$

and h , see (7), are concave, and moreover $h(o(\lambda(T))X^{(u)}(T)) \in L^2(\mathbb{R})$, $o(\lambda(T))h_x(o(\lambda(T))X^{(u)}(T)) \in L^2(\mathbb{R})$. If for $\bar{u} \in \mathcal{A}$ and the corresponding process $X^{(\bar{u})}$ satisfying equation (5) the following is true

$$H(t, X^{(\bar{u})}(t), \bar{u}(t), p(t), q(t)) - H(t, X^{(\bar{u})}(t), u, p(t), q(t)) \geq 0 \quad (13)$$

for every $t \in [0, T]$ and $u \in \mathbf{U}$ then the pair $\bar{u}, X^{(\bar{u})}$ is FMC.

Proof: since the proof follows rather standard technics we postpone it to Appendix.

□

4 Solution to control

The Hamiltonian for equation (2) and functional (3) takes the following form:

$$H(\omega, t, x, \pi, p, q) = p \left[rx - (\mu(t) + m(\hat{\lambda}(\omega, t) - r\mathbf{1}_n)\pi) \right] + \text{tr}(\sigma q^\top \pi) - e^{\delta t} C(t, \pi).$$

We assume that:

1. (to use theorem (3.1)) H is concave and U is concave as well. It happens like that if C is convex (growing marginal dissatisfaction from diverging from the required/desired level of risk) and if U is a utility function.
2. (to obtain π^* from relation (13)) C'_π is monotone in π .

Thus, the candidate for optimal control π^* can be written down in terms of p and q satisfying adjoint equation (11) from *Stochastic Maximum Principle*

$$\pi_i^*(t) = I^{-1} \left(t, (\mu_i(t) + m(\hat{\lambda}(t)) - r\mathbf{1}_n)p_i(t) + \sigma_{i,\cdot}(t) \cdot q(t)^\top \right).$$

We have to study the existence of solutions to adjoint equation. Firstly, we will consider FBSDEs with general forward equation.

The stochastic maximum principle leads to the following FBSDE (with coefficients of X depending on the choice of cost function C) with $\zeta_T := o(\lambda(T))$:

$$\begin{cases} dP(t) &= rP(t)dt + Q(t)dW(t) + Q_\lambda(t)d\hat{\lambda}(t) \\ P(T) &= \mathbf{E} [\zeta_T U_x(\zeta_T X(T)) | \mathcal{G}_T] \\ dX(t) &= l(t, X(t), P(t), Q(t))dt + \Sigma(t, X(t), P(t), Q(t))dW(t) \\ X(0) &= \xi, \end{cases} \quad (14)$$

We have to guarantee that the corresponding FBSDE has a solution. We use contraction mapping results (letting for simplicity $w := [W^\top \hat{\lambda}]^\top$).

Consider now the following system of FBSDEs:

$$\begin{cases} dY(t) &= H_x(X(t), Y(t), Z(t))dt + Z(t)dw(t) \\ Y(T) &= \mathbf{E} [\zeta_T h_x(\zeta_T X(T)) | \mathcal{G}_T] \\ dX(t) &= l(t, X(t), Y(t), Z(t))dt + \Sigma(t, X(t), Y(t), Z(t))dW(t) \\ X(0) &= \xi, \end{cases} \quad (15)$$

with random variable $\zeta_T \in L^\infty$, functions l and Σ Lipschitz with respect to variable x and satisfying the following growth conditions:

$$\begin{aligned} |l(x, y, z)| &\leq L_1|x| + L_1^{(y)}|y| + L_1^{(z)}|z| \\ |\Sigma(x, y, z)| &\leq K_1|x| + K_1^{(y)}|y| + K_1^{(z)}|z| \\ |\zeta_T h_x(\zeta_T x)| &\leq L^{(o)}|x|. \end{aligned}$$

Theorem 4.1 *Consider the above assumptions related to equation (15). Assume additionally that $\xi \in L^2_{\mathcal{F}}$ and H_x is Lipschitz continuous with respect to variables x , y and z with constants $L_H^{(x)}$, $L_H^{(y)}$ and $L_H^{(z)}$ respectively. If either 1 or 2 holds:*

1. $\hat{h}: \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ given as $\hat{h}(\omega, x) = \zeta_T(\omega)h_x(\zeta_T(\omega)x)$ is monotone in x ;
2. $H_x(x, y, z) := H_x^{(1)}(y, z)$ does not depend on x -variable and there exists $\beta > 0$ that

$$1 > 12 \left(L^{(0)}\right)^2 e^{(\beta+6(L_1^2+L_2^2))T} \left(L_1^{(z)} + K_1^{(z)}\right) =: 1 - \epsilon_L; \quad (16)$$

and

$$\beta > \frac{1}{2} \max\{L_H^{(y)}, L_H^{(z)}\} \quad (17)$$

then there exists a unique strong (adapted to \mathbb{G}) solution of (15) in $\mathcal{L}_T^\infty(\mathbb{R}, \mathbb{G}) \times \mathcal{L}_T^2(\mathbb{R}^{1 \times n}, \mathbb{G})$.

Remark 4.1 *The dependence of diffusion and drift on backward solutions y and z is required by the optimal solution which is a function of both processes.*

Remark 4.2 *Notice that condition (16) do not let the variance of Z to grow excessively — otherwise a highly volatile process Z would have to be applied to compensate the volatility of Y and $Z \notin \mathcal{L}_T^2$.*

Remark 4.3 *There might be useful to use nonmonotone \hat{h} in the case of modeling future investment opportunities for which there is a threshold (minimal) amount of funds needed x_{min} — for example minimal costs that has to be incurred to enter new market. In such a case „utility” U could have*

decreasing marginal increments below the value x_{\min} . But if the bank could accumulate $X^{(\pi)}(\omega, T)$ greater than x_{\min} then new investment possibilities could generate additional satisfaction (jump in satisfaction).

Proof: (theorem 4.1) In case of assumption 1 the proof is standard (see Peng and Wu (1999)). Let's move to the part with assumption 2. Let $g: = H_x$. We define firstly an auxiliary equation and show the existence of the solution.

$$\begin{cases} dY(t) &= g^{(n)}(t)dt + Z(t)dw(t) \\ Y(T) &= \mathbf{E} [\zeta_T h_x(\zeta_T X(T)) | \mathcal{G}_T] \\ dX(t) &= l(t, X(t), Y(t), z(t))dt + \Sigma(t, X(t), Y(t), z(t))dW(t) \\ X(0) &= \xi, \end{cases} \quad (18)$$

where $(\cdot)^{(n)}: = (\cdot \wedge n) \vee (-n)$ for the arguments $g \in \mathcal{L}_T^2(\mathbb{R}, \mathbb{G})$, and $z \in \mathcal{L}_T^2(\mathbb{R}^{1 \times n}, \mathbb{G})$. The definition of g_n is justified by further application.

The following result SDE with random coefficients is required (Yong and Zhou, 1999, Theorem 6.3).

Proposition 4.1 *Let $j_i(\cdot)$, $i \in \{1, 2\}$, be Lipschitz-continuous. The processes a_t , b_t belong to $\mathcal{L}_T^2(\mathbb{F})$ and $\xi \in L_{\mathcal{F}}^p$. Then the following SDE has unique strong solution on $[0, T]$*

$$\begin{cases} dx(t) &= (j_1(x(t)) + a_t) dt + (j_2(x(t)) + b_t) dW(t) \\ x(0) &= \xi \end{cases} \quad (19)$$

and $\sup_{t \in [0, T]} |x(t)|^p < K \mathbf{E} |\xi|^p$.

Let's note that in this simple (because of boundedness) the solution of the first two equations has the following form

$$Y^{(n)}(t) = \mathbf{E} \left[\zeta_T h_x^{(n)}(\zeta_T X(T)) - \int_t^T g^{(n)}(s) ds \middle| \mathcal{G}_t \right] \quad (20)$$

and is bounded. Since \mathbb{G} is \mathbf{P}_0 -Brownian filtration, the second process in the solution pair — $Z^{(n)}$ — is obtained by direct application of Martingale Representation Theorem and generalized Bayes formula to

$$\begin{aligned} M_t &= \mathbf{E} \left[\zeta_T h_x^{(n)}(\zeta_T X(T)) + \int_0^T g^{(n)}(s) ds \middle| \mathcal{G}_t \right] \\ &= \mathbf{E}_0 \left[\frac{\Lambda^{-1}(T)}{\mathbf{E}_0[\Lambda^{-1}(T) | \mathcal{G}_t]} \left(\zeta_T h_x^{(n)}(\zeta_T X(T)) + \int_0^T g^{(n)}(s) ds \right) \middle| \mathcal{G}_t \right]. \end{aligned}$$

To define contraction transformation we introduce the norm $\|\cdot\|_\beta$ in the Hilbert space $\mathcal{L}_T^2 \times \mathcal{L}_T^\infty \ni (y, z)$.

$$\|(y, z)\|_\beta = \mathbf{E} \left[\int_0^T e^{\beta s} (|y(s)|^2 + |z(s)|^2) ds \right].$$

Let's use the Itô Lemma to $(t, Y^{(n)}(t)) \rightarrow e^{\beta t} |Y^{(n)}(t)|^2$.

$$\begin{aligned} e^{\beta T} |Y^{(n)}(T)|^2 &= |Y^{(n)}(0)|^2 + \\ &\quad \beta \int_0^T e^{\beta s} |Y^{(n)}(s)|^2 ds + 2 \int_0^T e^{\beta s} Y^{(n)}(s) dY^{(n)}(s) + \int_0^T e^{\beta s} |Z^{(n)}(s)|^2 ds \end{aligned}$$

and

$$\begin{aligned} \beta \int_0^T e^{\beta s} |Y^{(n)}(s)|^2 ds + \int_0^T e^{\beta s} |Z^{(n)}(s)|^2 ds + |Y^{(n)}(0)|^2 &= \\ &= e^{\beta T} |Y^{(n)}(T)|^2 - 2 \int_0^T e^{\beta s} Y^{(n)}(s) dY^{(n)}(s) = \\ &= e^{\beta T} |\zeta_T h_x^{(n)}(\zeta_T X(T))|^2 + 2 \int_0^T e^{\beta s} |Y^{(n)}(s)| g^{(n)}(s) ds \\ &\quad + 2 \int_0^T e^{\beta s} |Y^{(n)}(s)| |Z^{(n)}(s)| dw(s) \end{aligned} \tag{21}$$

Notice that because of the particular form of equation for $X^{(n)}$ and truncation of $Y^{(n)}$ we have that $|X(\cdot)|^2 \in \mathcal{L}_T^2$. Then we take expectations of both

side of (21) and noting $\mathbf{E} \int_0^T e^{\beta s} |Y^{(n)}(s)| |Z^{(n)}(s)| dw(s) = 0$ we obtain

$$\begin{aligned}
& \mathbf{E} \left(\beta \int_0^T e^{\beta s} |Y^{(n)}(s)|^2 ds + \int_0^T e^{\beta s} |Z^{(n)}(s)|^2 ds \right) + \mathbf{E} |Y^{(n)}(0)|^2 \leq \\
& \leq \mathbf{E} e^{\beta T} |\zeta_T o^{(n)} X^{(n)}(T)|^2 + 2\mathbf{E} \int_0^T e^{\beta s} |Y^{(n)}(s)| g^{(n)}(s) ds \leq \\
& \leq \mathbf{E} e^{\beta T} |\zeta_T h_x^{(n)}(\zeta_T X(T))|^2 + \\
& + \mathbf{E} \int_0^T \frac{\beta}{2} e^{\beta s} |Y^{(n)}(s)|^2 ds + \mathbf{E} \int_0^T \frac{2}{\beta} |g^{(n)}(s)|^2 ds. \quad (22)
\end{aligned}$$

Applying Burkholder-Grundy-Davis inequality we get

$$\begin{aligned}
& \mathbf{E} e^{\beta T} |\zeta_T h_x^{(n)}(\zeta_T X(T))|^2 \leq \mathbf{E} e^{\beta T} |\zeta_T h_x(\zeta_T X(T))|^2 \leq \mathbf{E} e^{\beta T} |L^{(o)} X(T)|^2 \leq \\
& (L^{(o)})^2 \mathbf{E} e^{\beta T} \left| x_0 + \int_0^T l(t, X^{(n)}(t), Y^{(n)}(t), z(t)) dt + \right. \\
& \left. + \int_0^T \Sigma(t, X^{(n)}(t), Y^{(n)}(t), z(t)) dW(t) \right|^2 \leq \\
& \leq (L^{(o)})^2 e^{\beta T} \left(3|x_0|^2 + 6\mathbf{E} \int_0^T L_1^2 |X(s)|^2 ds + \right. \\
& + 6\mathbf{E} \int_0^T l_L^2(Y^{(n)}(s), z(s)) ds + 6\mathbf{E} \int_0^T K_1^2 |X^{(n)}(s)|^2 ds + \\
& \left. + 6\mathbf{E} \int_0^T \sigma_\Sigma^2(Y^{(n)}(s), z(s)) ds \right) \quad (23)
\end{aligned}$$

and also

$$v(t) \leq B + C \int_0^t v(s) ds \quad (24)$$

where

$$\begin{aligned}
v(t) &: = \mathbf{E} X^2(t), \\
B &: = 3|x_0|^2 + 6\mathbf{E} \int_0^T l_L^2(Y^{(n)}(s), z(s)) ds + 6\mathbf{E} \int_0^T \sigma_\Sigma^2(Y^{(n)}(s), z(s)) ds, \\
C &: = 6(L_1^2 + K_1^2).
\end{aligned}$$

But applying Gronwall lemma to inequality (24) we get from inequality (23) the following relation:

$$\mathbf{E}e^{\beta T}|\zeta_T h_x^{(n)}(\zeta_T X(T))|^2 \leq Be^{CT} \quad (25)$$

The inequality (16) implies that

$$\|(Y^{(n)}, Z^{(n)})\|_\beta^2 \leq \frac{2}{\beta} \mathbf{E} \int_0^T |g(s)|^2 ds. \quad (26)$$

Because of the convergence of RHS of the inequality we can proceed with n to ∞ and then for the Cauchy sequence $(Y^{(n)}, Z^{(n)})$ in a complete space $\mathcal{L}^2 \times \mathcal{L}^2$, we have the inequality for the solutions of the equation (18) with $g^{(n)}$ and $h_x^{(n)}$ replaced by the limits of $(g^{(n)})$ and $(h_x^{(n)})$ respectively.

The norms are equivalent for different β s and showing inequality (26) we concentrate on the one we have chosen and for which we have that

$$\|(Y^{(n)}, Z^{(n)})\|_\beta^2 \leq 4 \frac{(L_H^{(y)})^2}{\beta} \mathbf{E} \int_0^T e^{\beta t} |y(t)|^2 dt + 4 \frac{(L_H^{(z)})^2}{\beta} \mathbf{E} \int_0^T e^{\beta t} |z(t)|^2 dt$$

Thus, the transformation

$$\Psi : \mathcal{L}^2 \times \mathcal{L}^2 \rightarrow \mathcal{L}^2 \times \mathcal{L}^2$$

given by:

$$\begin{cases} dY(t) &= g(y(t), z(t))dt + Z(t)dw(t) \\ Y(T) &= \zeta_T o(X(T)) \\ dX(t) &= l(t, X(t), Y(t), z(t))dt + \Sigma(t, X(t), Y(t), z(t))dW(t) \\ X(0) &= \xi, \end{cases} \quad (27)$$

i.e.

$$\Psi(y, z) = (Y, Z).$$

is a contraction if such β satisfies inequality (17). To prove that $Y \in \mathcal{L}_T^\infty(\mathbb{R}, \mathbb{G})$ one can proceed exactly in line with Yong and Zhou (1999), inequality 2.30, making use of Lipschitz continuity of g .

□

Remark 4.4 *The only "tight" constraint is connected with the terminal condition for Y by incorporating of process Z in the forward equation — equation for X .*

5 Computationally tractable example

5.1 Parameters and helpful facts

Since the terminal wealth depends on unobserved λ it would be useful to transform that equation to more computable form. The special form of unobserved process allows to use the classical representation of nonlinear filter of $\mathbf{E}[U(o(\lambda(T))X^{(\pi)}(T))|\mathcal{G}(T)]$ in the form of SDE and then we will proceed with numerical approximation.

We assume that the unobserved measure of future market conditions λ is a random variable with known normal distribution with mean 0 and standard deviation S_0 . The observed $\hat{\lambda}$ process is a mean reversion Orstein-Uhlenbeck process.

$$\begin{aligned} d\hat{\lambda}(t) &= \alpha \left(\lambda(t) - \hat{\lambda}(t) \right) dt + \sigma_\lambda dW_\lambda^{(2)}(t), \quad d\lambda(t) = 0. \\ &\text{and } \lambda(0) \text{ is } \mathcal{N}(0, S_0) \end{aligned}$$

$$d\bar{\lambda}(t) = S^2(t) \frac{\alpha}{\sigma_\lambda} d\bar{W}(t) \quad \text{with} \quad S^2(t) = \frac{1}{\frac{1}{S_0} + \left(\frac{\alpha}{\sigma_\lambda}\right)^2 t} \quad (28)$$

and \hat{W} is the *innovation process* given by

$$d\bar{W}(t) = d\hat{\lambda}(t) - \alpha \left(\bar{\lambda}(t) - \hat{\lambda}(t) \right) dt.$$

On the other hand the dynamics of $\bar{\lambda}$ is given by

$$d\bar{\lambda}(t) = -\frac{S^2(t)\alpha^2}{\sigma_\lambda} \left(\bar{\lambda}(t) - \hat{\lambda}(t) \right) dt + S^2(t) \frac{\alpha}{\sigma_\lambda} d\hat{\lambda}(t)$$

with $(\mathbb{G}, \mathbf{P}_0)$ -Brownian motion. But it is easier to see from (28) how the bank *learns* about the true λ — process $\hat{\lambda}$ stabilizes on its historical mean since S tends to 0 in time.

We use the following standard fact.

Lemma 5.1 *For random variables Y and Z s.t. Y is \mathcal{H} -measurable and Borel function s.t. $\varphi(YZ) \in L^1(\mathcal{H})$ the following identity holds:*

$$\mathbb{E}[\varphi(Y, Z)|\mathcal{H}] = \mathbb{E}[\varphi(y, Z)|\mathcal{H}]|_{y=Y}.$$

Since we transform a problem of calculating a filter of a function of two arguments into a sequence of firstly filtering with fixed deterministic argument and then replacing it with the measurable component we have to compute conditional expectation of a deterministic function of a random variable. In general it could be troublesome nonlinear filtering exercise. In our case, strongly relying on the fact that $\mathbf{E}[\lambda|\mathcal{G}(t)]$ is Gaussian this can be done in the following way.

Lemma 5.2 *Let $o \in C^\infty(\mathbb{R})$. For any $(x_\lambda, x) \in \mathbb{R}^2$ define*

$$\Lambda(x_\lambda, x) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi S(T)}} o(y) U'(o(y)x) \exp \left\{ -\frac{(y - x_\lambda)^2}{2S(T)} \right\} dy. \quad (29)$$

Then

$$\mathbf{E} [o(\lambda) U'(o(\lambda)x) | \mathcal{G}(T)] = \Lambda(\bar{\lambda}(T), x) \quad (30)$$

with $\bar{\lambda}$ and S (solution of Riccati equation) given by (28).

Proof: Standard, see for example Bensoussan (1992) or Carmona and Ludkovsky (2004) since $U(h(\lambda)x) \in L^1(\mathbb{F})$. □

However in our example we take the very simple U — identical function $U(a) = a$. We are forced to do this to *get rid* of X from terminal condition in adjoint equation. In such a particular case adjoint equation becomes dependent on forward equations only by means of prediction of λ . It is independent of X and FBSDE becomes fully decoupled. It is necessary for application of numerical scheme proposed by Zhang (2004).

The particular values of basic parameters of the model are given in table (A).

5.2 Cost function – divergence from accepted risk level

In case of cost function C we take

$$C(t, \pi) = \alpha_1 ((\kappa - R(\pi))^+)^2 + \alpha_2 ((R(\pi) - \kappa)^+)^2.$$

If $\alpha_1 < \alpha_2$ then the cost C penalizes relatively more for investing in too risky way than in too secure way. The function R is the measure of the risk of

portfolio, e.g. it can be given as $R(\pi) = \sum_j (\sum_i |\sigma_{i,j}| \pi_i) / \sum_{i,j} |\sigma_{i,j}|$. In the numerical example we consider only one risky asset and we set $\alpha_1 = 0.1$ and $\alpha_2 = 0.5$. We will consider two different cases described by κ : a) a restrictive case with $\kappa = 0.3$; b) allowing for more risk in the assets with $\kappa = 0.6$. Thus, the cost has the form: $C(t, \pi) = 0.1 ((\kappa - \pi)^+)^2 + 0.5 ((\pi - \kappa)^+)^2$.

We assume that the economic condition $\hat{\lambda}$ influences the return from stocks through the function $m(l) = A \arctan(l)$, where A is a constant measuring the sensitivity of returns to variability of $\hat{\lambda}$. This is done only because of boundedness of \arctan and almost linearity around $l = 0$. We will illustrate the influence of $\hat{\lambda}$ to stock prices (and returns) by considering two versions of the model: with $A = 0.05$ (less sensitive process to the beliefs of future market conditions) and $A = 0.2$ (more sensitivity of S_i s to changes in $\hat{\lambda}$).

[Table 1 about here.]

5.3 Numerical procedure and convergence to optimality

The problem of solving equation (14) *explicit* can be gone round by using numerical procedure. Then the other problem arises — how far from the optimal value are the values obtained by substituting the optimal controls by their approximations? We will deal with this in the second part of this subsection.

We follow the scheme proposed by Zhang (2004). The n th nodes of time \mathbb{T}_n (called "nth partition") will be the sequence (t_0, t_1, \dots, t_n) s.t. $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$. We denote by $u_j^{(i)}(\mathcal{X}) = u^{(i)}(t_j, \mathcal{X})$ the value of

i th approximation function (i th partition) $u^{(i)}: \mathbb{T}_i \times \mathbb{R} \rightarrow \mathbb{R}$ at time t_j and in the state $(\hat{\varkappa})$. Additionally, we define one-step-forward approximation of X defined in equation (15), for state (x, \varkappa) and $j < i$.

$$\begin{aligned} X_{j+1}^{(i)}(x, \varkappa) &= x + l \left(t_j, x, u_{j+1}^{(i)}(\varkappa), v_{j+1}^{(i)}(\varkappa) \right) (t_{j+1} - t_j) + \\ &\quad \Sigma \left(t_j, x, u_{j+1}^{(i)}(\varkappa), v_{j+1}^{(i)}(\varkappa) \right) (W(t_{j+1}) - W(t_j)). \end{aligned} \quad (31)$$

Practically, x is equal to $X_j^{(i)}(x)$ and \varkappa is the realization of $\bar{\lambda}$ at time t_j . Let $\bar{\lambda}_j^{(i)}$ be a discretization of process $\bar{\lambda}$ in a stochastic Euler scheme and (analogously to $X_{j+1}^{(i)}$) let $\bar{\lambda}_j^{(i)}(\varkappa)$ be one-step-forward approximation. Given that (15) is independent of X (fully decoupled) and $o \in L^\infty(\mathbb{R})$ the numerical scheme for Y and Z for the n th partition is given in the following recursive way for the time axis nodes:

$$\begin{aligned} u_n^{(i)}(\varkappa) &= o(\varkappa) \\ U_{j+1}^{(i)}(\varkappa) &= u_{j+1}^{(i)} \left(\bar{\lambda}_j^{(i)}(\varkappa) \right) + \\ &\quad -g \left(t_j, u_{j+1}^{(i)}(\bar{\lambda}_j^{(i)}(\varkappa)), v_{j+1}^{(i)}(\bar{\lambda}_j^{(i)}(\varkappa)) \right) (t_{j+1} - t_j) \\ u_j^{(i)}(\varkappa) &= \mathbf{E} U_{j+1}^{(i)}(\varkappa) \end{aligned} \quad (32)$$

$$v_j^{(i)}(\varkappa) = \mathbf{E} \frac{1}{t_{j+1} - t_j} U_{j+1}^{(i)}(\varkappa) (W(t_{j+1}) - W(t_j)). \quad (33)$$

We get the following estimate

Theorem 5.1 *The numerical approximation of equation (15) given by*

$$\bar{\lambda}_{num}^{(i)}(t_j) = \bar{\lambda}_j^{(i)}(\bar{\lambda}_{num}^{(i)}(t_{j-1})), \quad \bar{\lambda}_{num}^{(i)}(0) = 0 \quad (34)$$

$$X_{num}^{(i)}(t_j) = X_j^{(i)}(X_{num}^{(i)}(t_{j-1}), \bar{\lambda}_{num}^{(i)}(t_{j-1})), \quad X_{num}^{(i)}(0) = x_0, \quad (35)$$

$$Y_{num}^{(i)}(t_j) = u_j^{(i)}(\bar{\lambda}_{num}^{(i)}(t_{j-1})) \quad (36)$$

$$Z_{num}^{(i)}(t_j) = v_j^{(i)}(\bar{\lambda}_{num}^{(i)}(t_{j-1})) \quad (37)$$

converges to the exact solution (X, Y, Z) in the following sense. The solution to the BSDE part satisfies:

$$\mathbf{E} \left[\int_0^T |Y(t) - Y^{(i)}(t)|^2 dt + \int_0^T |Z(t) - Z^{(i)}(t)|^2 dt \right] \leq C_{BSDE}(1 + |x_0|^2) |\mathbb{T}_i| \quad (38)$$

where

$$Y^{(i)}(t) = \sum_{j \in \{0, 1, \dots, N-1\}} \mathbf{1}_{t \in (t_j, t_{j+1}]} Y_{num}^{(i)}(t_j),$$

and analogously for Z (and also for X and $\bar{\lambda}$). For the forward part the following convergence holds:

$$X^{(i)}(\cdot) \xrightarrow{\mathcal{L}_T^2(\mathbb{R}, \mathbb{G})} X^*(\cdot). \quad (39)$$

Proof: The convergence of the scheme for P and Q given in relation (38) follows from Zhang (2004), theorem 5.3.

To prove the convergence in (39) we will formulate an auxiliary lemma. For a given norm \bar{m} , we introduce the following Hilbert space \mathcal{H} of $\{\mathcal{F}_t\}_{t \in [0, T]}$ -adapted processes with the norm

$$\|u(\cdot)\|_{\mathcal{H}} = \mathbf{E} \int_0^T (\bar{m}(|u(t)|))^2 dt.$$

It should be mentioned that the functions $\bar{m}(|u(\cdot)|)$ are from \mathcal{L}^2 .

Let's consider a process $\{y(t)\}_{t \in [0, T]}$ given by an equation

$$dY(t) = \bar{a}(t, y(t), u(t))dt + \bar{b}(t, y(t), u(t))dW(t) \quad (40)$$

where \bar{a} and \bar{b} satisfy conditions from Definition (3.1). The relation (39) is the consequence of the following lemma.

Lemma 5.3 *Let's take $\bar{u} \in \mathcal{H}$. Then for every sequence $(u_n) \subset \mathcal{H}$ s.t. $u_n \xrightarrow{\|\cdot\|_{\mathcal{H}}} \bar{u}$ the solutions y_n of equation (40) with u replaced by u_n converge to the solution \bar{y} with u replaced by \bar{u} with respect to $\|\cdot\|_{\mathcal{L}^2}$ e.i.*

$$y_n \xrightarrow{\|\cdot\|_{\mathcal{L}^2}} \bar{y}.$$

Proof: From the enhanced Lipschitz continuity condition, Burkholder-Davis-Grundy inequality and simple $(x_1 + x_2)^2 \leq 2x_1^2 + 2x_2^2$ inequality

$$\begin{aligned} \mathbf{E} |y(T) - y'(T)|^2 &\leq 2\mathbf{E} \left[\int_0^T [\bar{a}(t, y(t), u(t)) - \bar{a}(t, y'(t), u'(t))] dt \right]^2 + \\ &\quad 2\mathbf{E} \left[\int_0^T [\bar{b}(t, y(t), u(t)) - \bar{b}(t, y'(t), u'(t))] dW(t) \right]^2 \\ &\leq 4\mathbf{E} \left[\int_0^T (L_a |y(t) - y'(t)|^2 + (\bar{m}_a(|u(t) - u'(t)|))^2) dt \right] + \\ &\quad 4\mathbf{E} \left[\int_0^T (L_b |y(t) - y'(t)|^2 + (\bar{m}_b(|u(t) - u'(t)|))^2) dt \right] \\ &\leq 4\mathbf{E} \int_0^T (L_a + L_b) |y(t) - y'(t)|^2 dt + V(T) \end{aligned} \quad (41)$$

where

$$V(t) = \mathbf{E} \int_0^t [\bar{m}_a(|u(s) - u'(s)|)^2 + \bar{m}_b(|u(s) - u'(s)|)^2] ds$$

which is well-defined for $V(T) < \infty$ and $V(\cdot)$ is increasing. Hence, $U(\cdot)$ is integrable and (41) satisfies assumptions of Gronwall inequality, for $v(t) = \mathbf{E} |y(t) - y'(t)|^2$. So

$$\mathbf{E} |y(T) - y'(T)|^2 \leq 4V(0) + 16 \int_0^T e^{4(T-t)} V(t) dt. \quad (42)$$

Hence, the proof of theorem 5.1 is completed. □

Interpretation of results. The results of application of the numerical scheme to the model with particular parameters defined in subsection 5.1 are shown on figures (1), (2) and (3). In all cases the penalizing cost of investing in risky assets prevents from (or suggest to) keep riskiness of assets below κ . π^* is frequently much smaller then the threshold and if π^* exceeds κ then the difference is not huge. The optimal terminal wealth seems to be at least experimentally more variable if the bank is allowed to invest in more risky assets, i.e. if $\kappa = 0.6$. But also the mean terminal wealth is higher and the lower (experimental) percentile is not significantly smaller. It can be related to the higher return potential on the risky asset market comparing to risk-free bonds (r) and volatility σ .

[Figure 1 about here.]

[Figure 2 about here.]

It is not surprising that variability of optimal $X^*(T)$ is even higher if κ remains on the level of 0.6 and the variance of returns from risky asset increases, i.e. A changes from 0.05 to 0.2, see figures (2) and (3). On the other hand, it seems that if uncertainty of asset returns increases than the optimally investing bank may not necessarily gain in wealth at time T comparing to initial wealth x_0 .

[Figure 3 about here.]

One aspect of the application of numerical scheme requires special comment and further research. The resulting approximation of q is very rough and unstable. Instability means that if the simulation is repeated than the approximation in the same node (t_i, λ_i) may change by more than 100%. However the

range of changes remain similar. It means that the proposed method can directly be used to at least estimate the statistics of optimal wealth (moments). It may be doubtful to use it for calculation of exact optimal investment path. Some technics of variance reduction should have to be used before.

6 Comments and remarks

The optimization problem solved in the paper can pertain to each financial institution that faces some risk regulations. Also insurance companies must keep their capital on adequate level while being as much profitable as possible. Although regulations of pension and investment funds are less strict then in banking still capital adequacy is under scrutiny of regulators. The proposed model can be equally successfully applied to their case.

In practice, the accepted level of risk should refer to amount of capital possessed by a bank. Legal capital requirement like famous 8% rule is based on relation of risk-weighted assets and banks capital. Thus a natural extension of the model from our paper could deal with a cost function depending on π and X . The resulting FBSDE is much more difficult to solve.

There is a question how to check that optimal π^* and X^* related to it have *good* properties, i.e. whether for each t (Lebesgue measure a.s.!) $\sum_{i \in \{1, \dots, n\}} \pi_i^*(t) \leq X^*(t)$ (\mathbf{P} -a.s.) and $X^*(t) \geq 0$ \mathbf{P} -a.s. It means that we do not want a bank to have negative value of assets and investing more than it possesses.

The application of the numerical scheme which is proposed is strongly limited by the requirement to have decoupled forward and backward adjoint

equation. There is a scheme proposed by Delarue and Menozzi (2005) that treats the coupled FBSDEs but diffusion coefficient of forward equation is a function only of t , forward component X and backward Y . Dependence on Z is excluded and thus it can not be applied to portfolio optimization where diffusion is a function of portfolio process that is a function of both Y and Z . The success of *probabilistic* methods like that proposed by Zhang (2004) in portfolio choice strongly depends on overcoming this problem.

The natural extension or modification of the model could involve the dependence of asset prices and terminal wealth only on one unobserved process λ that could be learnt *only* by observation of asset prices. It seems to us that this is an open question how to make such a model tractable.

A Details of the proof of Theorem 3.1

Convexity of H implies that (the variable η_T is so far irrelevant)

$$\begin{aligned} & \int_0^T (H_x(t, \bar{X}(t), \bar{u}(t), p(t), q(t)))^\top (X^u(t) - \bar{X}(t)) dt \leq \\ & \int_0^T (H(t, X^u(t), u(t), p(t), q(t)) - H(t, \bar{X}(t), \bar{u}(t), p(t), q(t))) dt. \end{aligned} \quad (43)$$

Hence, for $\theta(t) =: X^u(t) - \bar{X}(t)$ and for any $u \in \mathcal{A}$ the following inequality is satisfied

$$E[p^\top(T)\theta(T)] \leq -E \int_0^T (f(t, X^u(t), u(t)) - f(t, \bar{X}(t), \bar{u}(t))) dt. \quad (44)$$

It is true because $\{\theta(t)\}_{t \in [0, T]}$ with $\theta(t) = X^u(t) - \bar{X}(t)$ satisfies equation:

$$\begin{aligned} d\theta(t) &= \left(b(t, X^{(u)}(t), \hat{\lambda}(t), u(t)) - b(t, \bar{X}(t), \hat{\lambda}(t), \bar{u}(t)) \right) dt + \\ &+ (\Sigma(t, X^{(u)}(t), u(t)) - \Sigma(t, \bar{X}(t), \bar{u}(t))) dW(t). \end{aligned} \quad (45)$$

We show what kind of SDE is satisfied by the process $\{p(t)\theta(t)\}_{t \in [0, T]}$. We use Itô Lemma for functions $J: \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$, $J(p, \theta) = p^\top \theta$ and additionally $\theta(t) = (\theta_1(t), \dots, \theta_k(t))$, $p(t) = (p_1(t), \dots, p_k(t))$ and

$$q(t) = \begin{bmatrix} q_{1,1}(t) & \cdots & q_{1,n}(t) \\ \vdots & \ddots & \vdots \\ q_{k,1}(t) & \cdots & q_{k,n}(t) \end{bmatrix}.$$

For a matrix A let $A^{(\cdot, i)}$ denote i th column and $A^{(i, \cdot)}$ – i th row. We have namely the following relation:

$$\begin{aligned} p(T)^\top \theta(T) &= p(0)^\top \theta(0) + \sum_{i=1}^k \int_0^T \theta_i(t) dp_i(t) + \sum_{i=1}^k \int_0^T p_i(t) d\theta_i(t) + \\ &\quad + \sum_{i=1}^k \int_0^T q_{\cdot, i}^\top(t) \left(\Sigma^{(\cdot, i)}(t, X^{(u)}(t), u(t)) - \Sigma^{(\cdot, i)}(t, \bar{X}(t), \bar{u}(t)) \right) dt \\ &= \sum_{i=1}^k \int_0^T -\theta_i(t) H_x^{(i)}(t, \bar{X}(t), \bar{u}(t), p(t), q(t)) dt \\ &\quad + \underbrace{\sum_{i=1}^k \int_0^T \theta_i(t) q_{i, \cdot}(t) dW(t)}_{(*)} \\ &\quad + \sum_{i=1}^k \int_0^T p_i(t) \left(b^{(i)}(t, X^{(u)}(t), \hat{\lambda}(t), u(t)) - b^{(i)}(t, \bar{X}(t), \hat{\lambda}(t), \bar{u}(t)) \right) dt \\ &\quad + \underbrace{\sum_{i=1}^k \int_0^T p_i(t) \left(\Sigma^{(i, \cdot)}(t, X^{(u)}(t), u(t)) - \Sigma^{(i, \cdot)}(t, \bar{X}(t), \bar{u}(t)) \right) dW(t)}_{(**)} \\ &\quad + \sum_{i=1}^k \int_0^T q_{\cdot, i}^\top(t) \left(\Sigma^{(\cdot, i)}(t, X^{(u)}(t), u(t)) - \Sigma^{(\cdot, i)}(t, \bar{X}(t), \bar{u}(t)) \right) dt. \quad (46) \end{aligned}$$

Because the functions b and σ satisfy extended Lipschitz condition (see Definition (3.1)) from the estimation of supremums of moments of diffusion process

with sufficiently regular drift and diffusion coefficients (Yong and Zhou, 1999, Chapter 1, theorem 6.16) the process $\{\theta(t)\}_{t \in [0, T]}$ belongs to $\mathcal{L}_T^2(\mathbb{R}^k)$, since there exists a constant $K_T \in \mathbb{R}_+$, such that

$$\mathbf{E} \max_{0 \leq t \leq T} |\theta(t)|^2 \leq K_T(1 + \mathbf{E}|\theta(0)|^2)$$

implies

$$\mathbf{E} \int_0^T |\theta(t)|^2 dt < \infty.$$

Similarly, the processes $\{p(t)\}_{t \in [0, T]}$ and $\{q(t)\}_{t \in [0, T]}$ belong to $\mathcal{L}_T^2(\mathbb{R}^k)$ and $\mathcal{L}_T^2(\mathbb{R}^{k \times n})$ respectively and it ensues from definition². Hence, stochastic integrals (*) and (**) can be treated as isometric integrals and therefore they are martingales (with mean equal to 0!) and acting with expectation operator on both sides of equation (46) and using inequality (43) lead to (44), because:

$$\begin{aligned} \mathbf{E} p(T)^\top \theta(T) &= E \left[\int_0^T -\theta^\top(t) H_x(t, \bar{X}(t), \bar{u}(t), p(t), q(t)) dt \right] \\ &\quad + \mathbf{E} \left[\int_0^T p(t)^\top \left(b(t, X^{(u)}(t), \hat{\lambda}(t), u(t)) - b(t, \bar{X}(t), \hat{\lambda}(t), \bar{u}(t)) \right) dt \right] \\ &\quad + \mathbf{E} \left[\int_0^T \text{tr} \left(q^\top(t) \left(\Sigma(t, X^{(u)}(t), u(t)) - \Sigma(t, \bar{X}(t), \bar{u}(t)) \right) \right) dt \right] \\ &\leq \mathbf{E} \int_0^T H(t, X^{(u)}(t), u(t), p(t), q(t)) - H(t, \bar{X}(t), \bar{u}(t), p(t), q(t)) dt \\ &\quad + \mathbf{E} \left[\int_0^T p(t)^\top \left(b(t, X^{(u)}(t), \hat{\lambda}(t), u(t)) - b(t, \bar{X}(t), \hat{\lambda}(t), \bar{u}(t)) \right) dt \right] \\ &\quad + \mathbf{E} \left[\int_0^T \text{tr} \left(q^\top(t) \left(\Sigma(t, X^{(u)}(t), u(t)) - \Sigma(t, \bar{X}(t), \bar{u}(t)) \right) \right) dt \right] \\ &= -\mathbf{E} \int_0^T \left(f(t, X^{(u)}(t), u(t)) - f(t, \bar{X}(t), \bar{u}(t)) \right) dt. \end{aligned} \quad (47)$$

²The process q lies in the subspace $\mathcal{L}_T^\infty(\mathbb{R}^{k \times n})$.

On the other hand, because h is concave, the following is true:

$$\begin{aligned} & \mathbf{E} \left(h(o(\lambda(T))X^{(u)}(T)) - h(o(\lambda(T))\bar{X}(T)) \right) = \\ & \mathbf{E} \left\{ \mathbf{E} \left[\left(h(o(\lambda(T))X^{(u)}(T)) - h(o(\lambda(T))\bar{X}(T)) \right) \middle| \mathcal{G}_T \right] \right\} \stackrel{(\diamond)}{\leq} \\ & \mathbf{E} \left[\mathbf{E} \left[o(\lambda(T))h_x(o(\lambda(T))\bar{X}(T))^\top \theta(T) \middle| \mathcal{G}_T \right] \right] = \mathbf{E}[p^\top(T)\theta(T)] \end{aligned} \quad (48)$$

(\diamond) follows from the concavity of h implying for each pair of random variables X_1, X_2 that $h(X_1) \leq h_x(X_1)^\top (X_1 - X_2)$. Thus from definition of conditional expectation, for each $A \in \mathcal{G}_T$ $\int_A h(X_1)d\mathbf{P} \leq \int_A h_x(X_1)^\top (X_1 - X_2)d\mathbf{P}$.

Combining inequalities (44) and (48) the following relation is obtained:

$$\begin{aligned} \mathbf{E} \left(h(\lambda(T)X^{(u)}(T)) - h(\lambda(T)\bar{X}(T)) \right) & \leq -\mathbf{E} \int_0^T \left(f(t, X^{(u)}(t), u(t)) - f(t, \bar{X}(t), \bar{u}(t)) \right) dt \\ \mathbf{E} \left[\int_0^T f(t, X^{(u)}(t), u(t)) dt + h(X^{(u)}(T)) \right] & \leq \mathbf{E} \left[\int_0^T f(t, \bar{X}(t), \bar{u}(t)) dt + h(\bar{X}(T)) \right]. \end{aligned}$$

u is arbitrary so the proof of theorem 3.1 is completed.

Acknowledgements

The paper was essentially developed during the stay in the King's College London supported by ESF program AMaMeF. Owing to financial support of National Bank of Poland the paper was presented on the First Conference of AMaMeF in April 26-29, 2006, Side, Antalya, Turkey. The author would like to thank Prof Lane Hughston, Prof Tom Hurd, Andrea Macrina and members of The Financial Mathematics Research Group in the King's College for valuable comments and suggestions, as well as Anna Karczewska for final corrections in the paper.

References

- Beněš V., Karatzas I., Ocone D., Wang H. (2004) *Control with Partial Observations and Explicit Solution to Mortensen's Equation*, Applied Mathematics & Optimization, 49, pp. 217–239.
- Basel Committee on Banking Supervision (1999) *A NEW CAPITAL ADEQUACY FRAMEWORK*, Consultative paper, Basel.
- Bensoussan A. (1992) *Stochastic Control of Partially Observable Systems* Cambridge University Press, Cambridge.
- Brendle S. (2005) *Portfolio Selection under Partial Observation and Filtering*, preliminary version of working paper, Princeton University.
- Cadenillas A., Karatzas I. (1995) *The stochastic maximum principle for linear, convex optimal control with random coefficients*. SIAM Journal on Control and Optimization, vol 33, number 2, pp. 590–624.
- Carmona R., Ludkovski M. (2004) *Convenience yield model with partial observations and exponential utility*, Princeton working paper.
- Cuoco D., Liu H. (2005) *An analysis of VaR-based capital requirements* Journal of Financial Intermediation (in press).
- Delarue F., Menozzi S. (2005) *A Forward-Backward Stochastic Algorithm for Quasi-Linear PDEs*, submitted to Annals of Applied Probability.
- Emmer S., Klüppelberg C., Korn R. (2001) *Optimal portfolios with bounded capital at risk*, Math. Finance 11, pp. 365–384.

- Lefèvre D. (2001) *An introduction to utility maximization with partial observation*, INRIA rapport de recherche, N° 4183.
- Karatzas I., Shreve S. E. (1997) *Brownian Motion and Stochastic Calculus* Springer-Verlag, New York.
- Kohlmann M., Zhou X. Y. (2000) *Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach*, SIAM J. Control Optim., 38, No. 5, pp. 1392–1407.
- Lipster R. Sh., Shiryaev A. N. (2001) *Statistics of Random Processes*. Springer (Berlin).
- Pástor Ľ., Veronesi P. (2003) *Stock Valuation and Learning about Profitability*, The Journal of Finance, Vol 58, pp. 1749–1789.
- Peng Sh., Wu Z. (1999) *Fully Coupled Forward-Backward Stochastic Differential Equations and Application to Optimal Control*, SIAM J. of Control&Optimization, Vol. 37, No. 3, pp. 825–843.
- Rivière O. (2005) *Equations différentielles stochastiques progressives rétrogrades: Equations aux dérivées partielles et discrétisation*, PhD Thesis, 1st August (in French).
- Santos, J. (2002) *Bank capital regulation in contemporary banking theory: A review of the literature*, Working Paper 90, Bank for International Settlements.
- Yong J., Zhou X. (1999) *Stochastic Controls. Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York.

Zhang J. (2004) *A Numerical Scheme for BSDEs*, The Annals of Applied Probability, Vol. 14, pp. 459–488.

Zhou X. Y (1998) *Stochastic Near-Optimal Controls: Necessary and Sufficient Conditions for Near-Optimality*, SIAM J. of Control & Optimization, Vol. 38, pp.929–947.

List of Figures

- 1 Comparison of results. The threshold $\kappa = 0.3$ and $A = 0.05$. . 37
- 2 Comparison of results. The threshold $\kappa = 0.6$ and $A = 0.05$. . 38
- 3 Comparison of results. The threshold $\kappa = 0.6$ and $A = 0.2$. . . 39

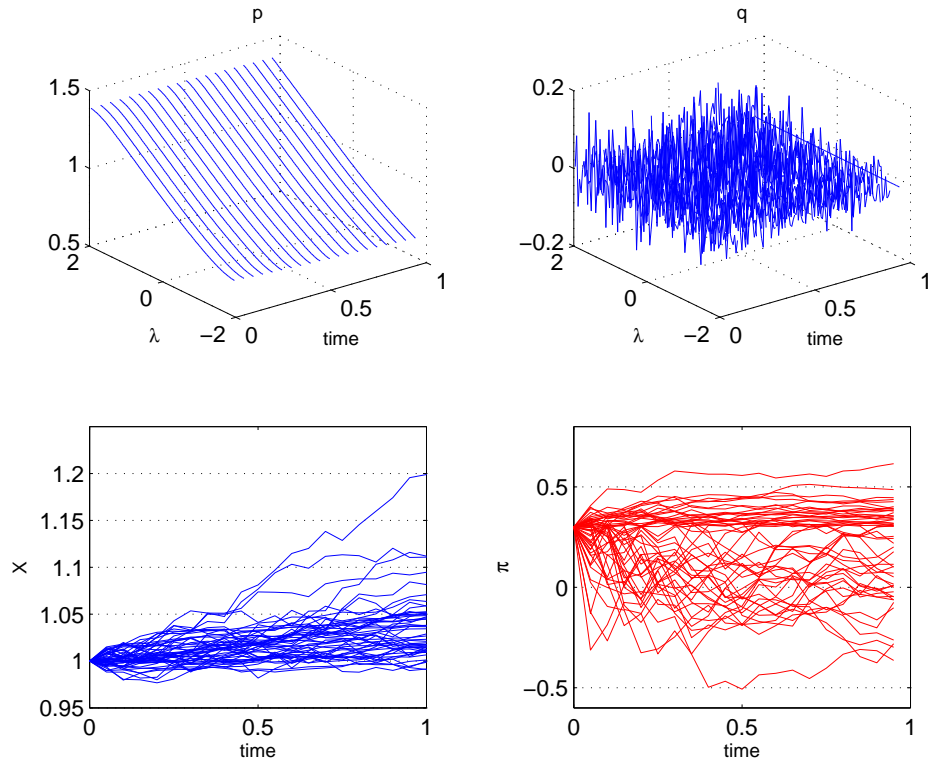


Figure 1: Comparison of results. The threshold $\kappa = 0.3$ and $A = 0.05$.

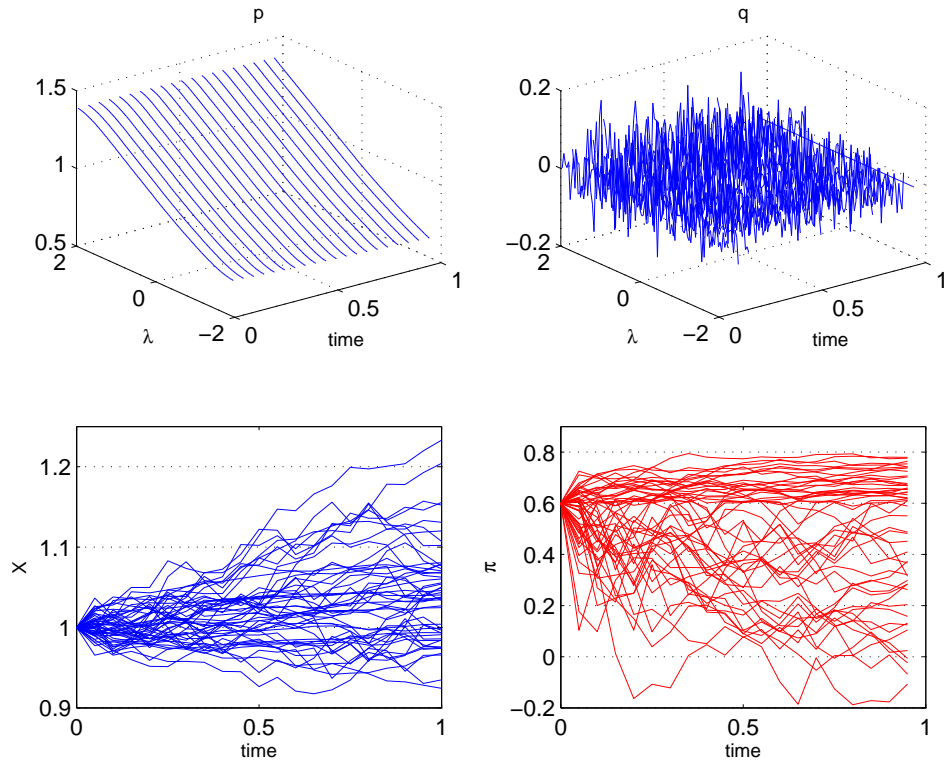


Figure 2: Comparison of results. The threshold $\kappa = 0.6$ and $A = 0.05$.

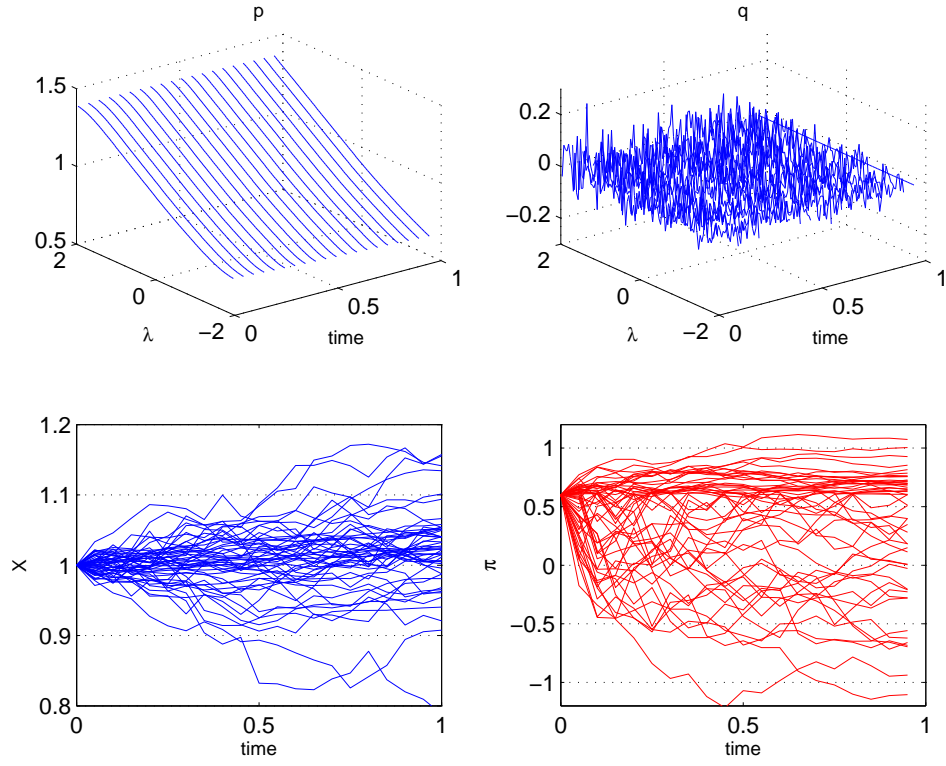


Figure 3: Comparison of results. The threshold $\kappa = 0.6$ and $A = 0.2$.

List of Tables

- 1 Summary of values of parameters assumed in the simulation. . 41

x_0	r	δ	μ_1	T	dt	σ	α	σ_λ	$S(0)$
1.00	0.04	1.00	0.08	1.00	0.05	0.10	0.20	0.10	1.00

Table 1: Summary of values of parameters assumed in the simulation.